

## UMR ADEF

### JOURNAL DU SEMINAIRE TAD/IDD

#### Théorie Anthropologique du Didactique & Ingénierie Didactique du Développement

*Toda pregunta implica la pérdida de una intimidad o el extinguirse de una adoración. María Zambrano (1904-1991)*

*There is a phrase I learned in college called, "having a healthy disregard for the impossible." That is a really good phrase. Larry Page (1973-?)*

*Le séminaire TAD & IDD est animé par Yves Chevallard au sein de l'équipe 1 de l'UMR ADEF, dont le domaine général de recherche s'intitule « École et anthropologie didactique des savoirs ».*

*Ce séminaire a, solidairement, une double visée. D'une part, il vise à mettre en débat des recherches (achevées, en cours ou en projet) touchant à la TAD ou, dans ce cadre, à des problèmes d'ingénierie didactique du développement, quel qu'en soit le cadre institutionnel. D'autre part, il vise à faire émerger les problèmes de tous ordres touchant au développement didactique des institutions, et notamment de la profession de professeur de mathématiques.*

*Deux domaines de recherche sont au cœur du séminaire : un domaine en émergence, la didactique de l'enquête codisciplinaire ; un domaine en devenir, la didactique des savoirs mathématiques.*

*La conduite des séances et leur suivi se fixent notamment pour objectif d'aider les participants à étendre et à approfondir leur connaissance théorique et leur maîtrise pratique de la TAD et des outils de divers ordres que cette théorie apporte ou permet d'élaborer.*

*Sauf exception, les séances se déroulent le vendredi après-midi, de 15 h à 17 h puis de 17 h 30 à 19 h 30, cette seconde partie pouvant être suivie en visioconférence.*

### → Séance 6 – Vendredi 2 mai 2008

## MATHÉMATIQUES DANS LA CITÉ

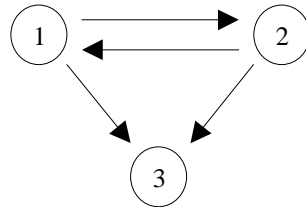
### 1. PageRank

a) Je change aujourd'hui l'ordre usuel des rubriques, et cela dans le but de donner plus de place à l'étude de l'algorithme PageRank de Google. On se rappelle que, pour attribuer un rang  $r_i$  à chacune des pages  $P_i$  du Web, Sergey Brin et Larry Page doivent résoudre l'équation  $\mathbf{R}^T = \mathbf{R}^T \mathbf{H}$ , où  $\mathbf{R}^T$  est le vecteur-ligne  $[r_1 \ r_2 \ \dots \ r_i \ \dots \ r_n]$  et où  $\mathbf{H}$  est la matrice dont l'élément  $H_{ij}$  situé à la  $i$ -ième ligne et à la  $j$ -ième colonne vaut 0 si la page  $P_i$  n'envoie pas de lien vers la page  $P_j$ , et  $\frac{1}{|P_i|}$  sinon, où  $|P_i|$  désigne le nombre de liens qui *partent* de  $P_i$  :

$$H_{ij} = \begin{cases} \frac{1}{|P_i|} & \text{s'il y a un lien de } P_i \text{ vers } P_j \\ 0 & \text{sinon} \end{cases} .$$

b) Pour résoudre l'équation  $\mathbf{R}^T = \mathbf{R}^T \mathbf{H}$ , on pourrait tenter de procéder *algébriquement*, comme on le fait couramment pour résoudre, au collège, une équation du 1<sup>er</sup> degré à une inconnue numérique  $x$ . Considérons par exemple l'équation  $x = 2x + 1$ . Dans ce cas on aurait :  $x = 2x + 1 \Leftrightarrow x - 2x = 1 \Leftrightarrow x(1 - 2) = 1 \Leftrightarrow x(1 - 2)(1 - 2)^{-1} = 1 \times (1 - 2)^{-1} \Leftrightarrow x \times 1 = (1 - 2)^{-1} \Leftrightarrow x = \frac{1}{1 - 2} = \frac{1}{-1} = -1$ . On a ainsi une *solution unique*.

► La même « manœuvre » sur l'équation à une inconnue *matricielle*  $\mathbf{R}^T = [r_1 \ r_2 \ \dots \ r_n]$  conduit d'abord à écrire ceci :  $\mathbf{R}^T = \mathbf{R}^T \mathbf{H} \Leftrightarrow \mathbf{R}^T - \mathbf{R}^T \mathbf{H} = \mathbf{0}$ . Afin de poursuivre, il me faut introduire la matrice *unité* d'ordre  $n$ , notée  $\mathbf{I}$  (ou  $\mathbf{I}_n$ ), dont tous les coefficients sont nuls, à l'exception des coefficients  $(i, i)$ , dits *diagonaux*, qui sont tous égaux à 1. Cette matrice correspond, pour ce qui est du calcul matriciel, au nombre 1 du calcul avec des nombres. On aura alors en effet :  $\mathbf{R}^T - \mathbf{R}^T \mathbf{H} = \mathbf{R}^T \mathbf{I} - \mathbf{R}^T \mathbf{H} = \mathbf{R}^T (\mathbf{I} - \mathbf{H})$ . Vérifions-le sur un exemple, celui où le Web considéré se présente ainsi.



La matrice  $\mathbf{H}$  est ici la matrice  $3 \times 3$  que voici :  $\mathbf{H} = \begin{bmatrix} 0 & 0,5 & 0,5 \\ 0,5 & 0 & 0,5 \\ 0 & 0 & 0 \end{bmatrix}$ . La matrice  $\mathbf{I}$

(=  $\mathbf{I}_3$ ) est celle-ci :  $\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ . Le vecteur des rangs s'écrit  $\mathbf{R}^T = [r_1 \ r_2 \ r_3]$ . On a

bien (on le vérifiera) :  $\mathbf{R}^T \mathbf{I} = [r_1 \ r_2 \ r_3] \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = [r_1 \ r_2 \ r_3] = \mathbf{R}^T$ . On admettra que,

par ailleurs, on a en général  $\mathbf{A} \mathbf{B} - \mathbf{A} \mathbf{C} = \mathbf{A} (\mathbf{B} - \mathbf{C})$ , où  $\mathbf{A}$  est une matrice  $1 \times n$  tandis que  $\mathbf{B}$  et  $\mathbf{C}$  sont des matrices  $n \times n$  (de sorte que la matrice  $\mathbf{B} - \mathbf{C}$  est elle-même  $n \times n$ ). Il vient donc ici :  $\mathbf{R}^T \mathbf{I} - \mathbf{R}^T \mathbf{H} = \mathbf{R}^T (\mathbf{I} - \mathbf{H})$ . On a en l'espèce :

$$\mathbf{I} - \mathbf{H} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & 0,5 & 0,5 \\ 0,5 & 0 & 0,5 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & -0,5 & -0,5 \\ -0,5 & 1 & -0,5 \\ 0 & 0 & 1 \end{bmatrix}$$

L'équation matricielle à résoudre, soit  $\mathbf{R}^T (\mathbf{I} - \mathbf{H}) = \mathbf{0}$ , s'écrit donc :

$$[r_1 \ r_2 \ r_3] \begin{bmatrix} 1 & -0,5 & -0,5 \\ -0,5 & 1 & -0,5 \\ 0 & 0 & 1 \end{bmatrix} = [0 \ 0 \ 0].$$

Il s'agit de déterminer s'il existe des nombres  $r_1, r_2, r_3$  vérifiant le système d'équations suivant :

$$\begin{cases} r_1 - 0,5 r_2 - 0,5 r_3 = 0 \\ -0,5 r_1 + r_2 - 0,5 r_3 = 0 \\ r_3 = 0 \end{cases} .$$

Bien entendu, il est clair qu'il existe toujours la solution *nulle*, qui s'écrit  $\mathbf{R}^T = [r_1 \ r_2 \ r_3] = \mathbf{0} = [0 \ 0 \ 0]$ , soit encore  $r_1 = r_2 = r_3 = 0$ . Ce qui nous intéresse, c'est de savoir s'il existe une solution *non nulle*, c'est-à-dire telle qu'au moins l'un des nombres  $r_1, r_2, r_3$  soit non nul. En fait, dans le cas qui nous occupe (définir le rang des pages  $P_i$ ), il faudrait même que l'on ait  $r_1, r_2, r_3 > 0$ .

► Il convient de noter que si, au lieu de matrices, nous avons affaire à des nombres, c'est-à-dire si nous devons résoudre en  $x$  une équation de la forme  $x \times k = 0$  où  $k$  est un nombre non nul (« correspondant » à la matrice non nulle  $\mathbf{I} - \mathbf{H}$ ), nous n'aurions qu'une seule solution :  $x = 0$ . Ce qu'on espère donc c'est que ce cas *ne se produira pas ici*.

► Dans le cas particulier examiné, on peut résoudre le système en  $(r_1, r_2, r_3)$  en éliminant d'abord  $r_3$  : puisque  $r_3 = 0$ , le système s'écrit alors

$$\begin{cases} r_1 - 0,5 r_2 = 0 \\ -0,5 r_1 + r_2 = 0 \end{cases}$$

La 2<sup>e</sup> équation donne  $r_2 = 0,5 r_1$  ; la 1<sup>re</sup> donne alors :  $r_1 - 0,25 r_1 = 0,75 r_1 = 0$ . Cela implique que  $r_1$ , et donc aussi  $r_2$ , est *nul*. Le système n'admet que la solution nulle.

► On peut vérifier que la matrice

$$\mathbf{I} - \mathbf{H} = \begin{bmatrix} 1 & -0,5 & -0,5 \\ -0,5 & 1 & -0,5 \\ 0 & 0 & 1 \end{bmatrix}$$

admet une matrice *inverse* (comme 2 admet le nombre inverse  $2^{-1} = \frac{1}{2}$ ), c'est-à-dire une matrice, que l'on notera  $(\mathbf{I} - \mathbf{H})^{-1}$ , telle que l'on ait :  $(\mathbf{I} - \mathbf{H})(\mathbf{I} - \mathbf{H})^{-1} = \mathbf{I}$ . On a ici en fait :

$$(\mathbf{I} - \mathbf{H})^{-1} = \begin{bmatrix} 4/3 & 2/3 & 1 \\ 2/3 & 4/3 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Il vient en effet :

$$\begin{bmatrix} 1 & -0,5 & -0,5 \\ -0,5 & 1 & -0,5 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 4/3 & 2/3 & 1 \\ 2/3 & 4/3 & 1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

► Pour déterminer, au moins dans les cas les plus simples, si une matrice admet une inverse, et quelle est cette matrice inverse, on peut recourir à un logiciel en ligne de calcul matriciel, tel celui que l'on trouvera à cette adresse : [http://wims.unice.fr/wims/en\\_tool~linear~matrix.html](http://wims.unice.fr/wims/en_tool~linear~matrix.html). Le fait qu'une matrice  $\mathbf{A}$  (ici,  $\mathbf{I} - \mathbf{H}$ ) est inversible se voit d'une part au fait que  $\text{rank}(\mathbf{A})$  (ce qui, en français, désigne le *rang* de  $\mathbf{A}$ ) est égal à l'ordre de la matrice (ici, 3) ou encore au fait que  $\det(\mathbf{A})$  (ce qui, en français, désigne le *déterminant* de  $\mathbf{A}$ ) est non nul. On vérifiera que la matrice  $\mathbf{A}^{-1}$  confirme le résultat établi plus haut.

**Matrix calculator**

$$A = \begin{pmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 & -\frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix}$$

rank(A) = 3  
det(A) = 3/4  
trace(A) = 3  
The matrix is not symmetric.

$$A^{-1} = \begin{pmatrix} \frac{4}{3} & \frac{2}{3} & 1 \\ \frac{2}{3} & \frac{4}{3} & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

► Lorsqu'une matrice  $(\mathbf{I} - \mathbf{H})^{-1}$  existe, on a :

$$\mathbf{R}^T(\mathbf{I} - \mathbf{H}) = \mathbf{0} \Leftrightarrow \mathbf{R}^T(\mathbf{I} - \mathbf{H})(\mathbf{I} - \mathbf{H})^{-1} = \mathbf{0}$$

et donc  $\mathbf{R}^T \mathbf{I} = \mathbf{0}$  ou  $\mathbf{R}^T = \mathbf{0}$ . En d'autres termes, lorsque la matrice  $\mathbf{I} - \mathbf{H}$  est *invertible* (= admet une matrice inverse), l'équation matricielle  $\mathbf{R}^T = \mathbf{R}^T \mathbf{H}$  n'admet que la solution nulle : le vecteur  $\mathbf{R}^T$  espéré *n'existe pas*.

c) Revenons maintenant à l'idée de rechercher une solution  $\mathbf{R}^T$  comme limite de la suite de vecteurs lignes  $\mathbf{R}^{(0)T}, \mathbf{R}^{(1)T}, \mathbf{R}^{(2)T}, \dots, \mathbf{R}^{(k)T}, \mathbf{R}^{(k+1)T}, \dots$ , définie par

$$\mathbf{R}^{(0)T} = \left[ \frac{1}{n} \frac{1}{n} \dots \frac{1}{n} \dots \frac{1}{n} \right]$$

(où  $n = |W|$ ) et par la relation de récurrence  $\mathbf{R}^{(k+1)T} = \mathbf{R}^{(k)T} \mathbf{H}$ .

► On notera d'abord que cette manière de tenter de résoudre l'équation existe aussi – même si l'on ne s'en sert pas – pour les humbles équations du premier degré du collège. Considérons à nouveau l'équation  $x = 2x + 1$ . Posons  $x_0 = 0$  ; la suite  $x_0, x_1, x_2, \dots, x_k, \dots$  définie par la relation de récurrence  $x_{k+1} = 2x_k + 1$  a les valeurs suivantes :  $x_1 = 2 \times 0 + 1 = 1$  ;  $x_2 = 2 \times 1 + 1 = 3$  ;  $x_3 = 2 \times 3 + 1 = 7$  ;  $x_4 = 2 \times 7 + 1 = 15$  ;  $x_5 = 2 \times 15 + 1 = 31$  ;  $x_6 = 2 \times 31 + 1 = 63$ , etc. On voit à l'évidence que la suite  $(x_k)$  ne convergera pas vers l'unique solution de l'équation proposée, soit  $x = -1$  : en fait, la suite « converge » *vers l'infini*.

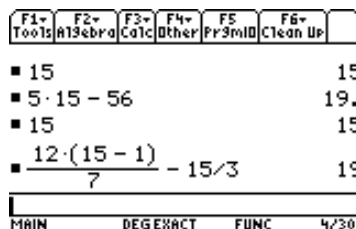
► On peut cependant « travailler » un peu l'équation pour lui donner une forme engendrant une suite convenable : au lieu de  $x = 2x + 1$ , écrivons-la, en ajoutant d'abord  $-1$  aux deux membres, puis en les divisant par 2, sous la forme

$$\frac{x-1}{2} = x$$





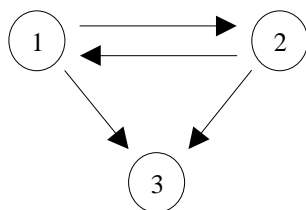
Il semble que la suite converge vers 15. De fait, on peut vérifier alors que les deux membres de l'équation initiale sont égaux, comme on le voit ci-après.



► Notons en passant que la résolution algébrique peut tirer parti du résultat obtenu. Posons en effet  $x = 15 + 21y$ ; il vient :  $5x - 56 = 75 + 105y - 56 = 105y + 19$  et  $\frac{12(x-1)}{7} - \frac{x}{3} = \frac{12(14 + 21y)}{7} - \frac{15 + 21y}{3} = 12(2 + 3y) - 5 - 7y = 29y + 19$ . L'équation donnée équivaut à  $105y + 19 = 29y + 19$ , soit à  $76y = 0$  ou  $y = 0$ . L'unique solution de l'équation en  $x$  est donc  $x = 15 + 21 \times 0 = 15$ .

d) Que se passe-t-il quand on opère par itérations sur l'équation  $\mathbf{R}^T = \mathbf{R}^T \mathbf{H}$  ?

► Reprenons le cas du Web à trois pages vu plus haut.



On a vu que, ici, on a :  $\mathbf{H} = \begin{bmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 0 & 0 & 0 \end{bmatrix}$ . On sait que, la matrice  $\mathbf{I} - \mathbf{H}$  étant inversible, le vecteur recherché n'existe pas.

► On prend  $\mathbf{R}^{(0)T} = [1/3 \ 1/3 \ 1/3]$ ; il vient :

$$\mathbf{R}^{(1)T} = \mathbf{R}^{(0)T} \mathbf{H} = [1/3 \ 1/3 \ 1/3] \begin{bmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 0 & 0 & 0 \end{bmatrix} = [1/6 \ 1/6 \ 1/3];$$

$$\mathbf{R}^{(2)T} = [1/6 \ 1/6 \ 1/3] \begin{bmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 0 & 0 & 0 \end{bmatrix} = [1/12 \ 1/12 \ 1/6];$$

$$\mathbf{R}^{(3)T} = [1/12 \ 1/12 \ 1/6] \begin{bmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 0 & 0 & 0 \end{bmatrix} = [1/24 \ 1/24 \ 1/12].$$

Il semble que, pour  $k \geq 1$ , on aurait plus généralement

$$\mathbf{R}^{(k)T} = [1/3 \cdot 2^{-k} \ 1/3 \cdot 2^{-k} \ 1/3 \cdot 2^{-(k-1)}].$$

La suite des vecteurs  $\mathbf{R}^{(k)T}$  tend donc vers le vecteur nul.

► Pour utiliser le logiciel en ligne déjà employé afin de vérifier les résultats obtenus, il convient d'utiliser un petit résultat algébrique. On a :

$$\mathbf{R}^{(2)T} = \mathbf{R}^{(1)T} \mathbf{H} = (\mathbf{R}^{(0)T} \mathbf{H}) \mathbf{H} = \mathbf{R}^{(0)T} \mathbf{H}^2;$$

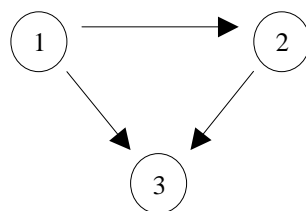
$$\mathbf{R}^{(3)T} = \mathbf{R}^{(2)T} \mathbf{H} = (\mathbf{R}^{(0)T} \mathbf{H}^2) \mathbf{H} = \mathbf{R}^{(0)T} \mathbf{H}^3;$$

D'une façon générale, on a :  $\mathbf{R}^{(k)T} = \mathbf{R}^{(0)T} \mathbf{H}^k$ . Cela noté, on obtient ceci (où  $\mathbf{M1} = \mathbf{H}$  et  $\mathbf{M2} = [1/3 \ 1/3 \ 1/3]$ ), qui confirme les résultats précédents.

$$\mathbf{M2} * \mathbf{M1}^3 = \left( \frac{1}{24} \quad \frac{1}{24} \quad \frac{1}{12} \right)$$

$$\mathbf{M2} * \mathbf{M1}^{10} = \left( \frac{1}{3072} \quad \frac{1}{3072} \quad \frac{1}{1536} \right)$$

► Considérons maintenant le Web à trois pages que voici.



On a ici :  $\mathbf{H} = \begin{bmatrix} 0 & 1/2 & 1/2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ . Il vient :

$\mathbf{I} - \mathbf{H} = \begin{bmatrix} 1 & -1/2 & -1/2 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$ . Le logiciel utilisé montre alors (ci-après) que le vecteur des rangs espéré n'existe pas.

**Matrix calculator**

$$A = \begin{pmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$$

rank(A) = 3  
det(A) = 1  
trace(A) = 3  
The matrix is not symmetric.

$$A^{-1} = \begin{pmatrix} 1 & \frac{1}{2} & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

► Que se passe-t-il pour ce qui est de l'itération ? Il vient :

$$\mathbf{R}^{(1)T} = \mathbf{R}^{(0)T} \mathbf{H} = [1/3 \ 1/3 \ 1/3] \begin{bmatrix} 0 & 1/2 & 1/2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = [0 \ 1/6 \ 1/2];$$

$$\mathbf{R}^{(2)T} = [0 \ 1/6 \ 1/2] \begin{bmatrix} 0 & 1/2 & 1/2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = [0 \ 0 \ 1/6];$$

$$\mathbf{R}^{(3)T} = [0 \ 0 \ 1/6] \begin{bmatrix} 0 & 1/2 & 1/2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = [0 \ 0 \ 0].$$

En fait on a ceci (où  $\mathbf{A} = \mathbf{H}$ ).

**Matrix calculator**

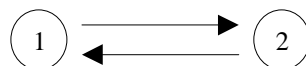
$$A = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

$$A^2 = \begin{pmatrix} 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$A^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Le fait que  $\mathbf{H}^3 = 0$  confirme le résultat antérieurement obtenu.

► Considérons le Web suivant.



On a ici pour matrice  $\mathbf{H} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$  et donc



$$\mathbf{I} - \mathbf{H} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

Ici, le logiciel utilisé indique que  $\mathbf{A} = \mathbf{I} - \mathbf{H}$  n'est pas inversible, ce qui laisse espérer qu'il soit possible de trouver un vecteur de rangs.

**Matrix calculator**

$A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$

rank(A) = 1  
det(A) = 0  
trace(A) = 2  
signature(A) = (1,0)

De fait on a :

$$[0,5 \ 0,5](\mathbf{I} - \mathbf{H}) = [0,5 \ 0,5] \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} = [0 \ 0] = \mathbf{0}.$$

► Que se passe-t-il quant à l'itération de la relation  $\mathbf{R}^{(1)T} = \mathbf{R}^{(0)T} \mathbf{H}$  ? Notons d'abord qu'on a ceci (où  $\mathbf{A} = \mathbf{H}$ ).

**Matrix calculator**

$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

$A^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$   
 $A^3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

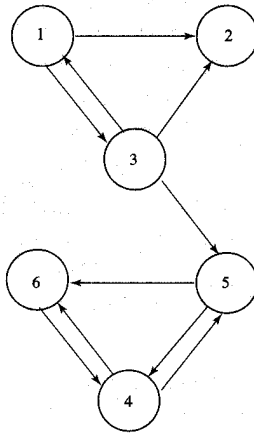
Plus généralement,  $\mathbf{H}^k = \mathbf{H}$  lorsque  $k$  est impair,  $= \mathbf{I}$  lorsque  $k$  est pair.

Lorsque  $k$  est impair, donc, on a :  $\mathbf{R}^{(k)T} = \mathbf{R}^{(0)T} \mathbf{H} = [0,5 \ 0,5] \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = [0,5 \ 0,5]$  ;

lorsque  $k$  est pair, on a :  $\mathbf{R}^{(k)T} = \mathbf{R}^{(0)T} \mathbf{I} = [0,5 \ 0,5] \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = [0,5 \ 0,5]$ . Ainsi, si  $\mathbf{R}^{(0)T} = [0,5 \ 0,5]$ , on a  $\mathbf{R}^{(k)T} = [0,5 \ 0,5]$  pour tout  $k$ .

► Prenons alors pour vecteur  $\mathbf{R}^{(0)T}$  le vecteur  $[1 \ 0]$ . Si  $k$  est impair, on aura  $\mathbf{R}^{(k)T} = \mathbf{R}^{(0)T} \mathbf{H} = [1 \ 0] \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = [0 \ 1]$  ; et, si  $k$  est pair :  $\mathbf{R}^{(k)T} = \mathbf{R}^{(0)T} \mathbf{I} = [1 \ 0] \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = [1 \ 0]$ . La suite des vecteurs  $\mathbf{R}^{(k)T}$  va osciller entre les vecteurs  $[1 \ 0]$  et  $[0 \ 1]$ , sans converger.

e) Les difficultés précédentes vont être dépassées par les "Google boys". Avant de voir comment, on considèrera un exemple déjà évoqué, celui du Web suivant.



► On voit ici que le sous-Web formé des pages  $P_4, P_5, P_6$  n'envoie aucun lien vers la partie formée des pages  $P_1, P_2, P_3$ . On peut s'attendre à ce que cette partie « thésaurise » (*hoard*) le rang au détriment de la partie  $\{P_1, P_2, P_3\}$ . Qu'en est-il ?

► La matrice  $\mathbf{H}$  est ici la suivante :

$$\mathbf{H} = \begin{bmatrix} 0 & 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1/3 & 1/3 & 0 & 0 & 1/3 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

Le logiciel de calcul matriciel déjà utilisé donne :

$$\mathbf{R}^{(13)T} = \mathbf{R}^{(0)T} \mathbf{H}^{13} = \begin{bmatrix} 1 & 5 & 1 & 3185267 & 14330741 & 55987 \\ 839808 & 1679616 & 559872 & 11943936 & 107495424 & 279936 \end{bmatrix}$$

On voit que les valeurs correspondant aux pages  $P_1, P_2, P_3$  sont quasiment nulles : tout le rang est accaparé par les pages  $P_4, P_5, P_6$ ; on a encore :

$$\mathbf{R}^{(13)T} \approx [0 \ 0 \ 0 \ 0,27 \ 0,13 \ 0,2].$$

(De façon énigmatique, Langville et Meyer donnent, page 35 de leur livre, des valeurs approchées quelque peu différentes :  $[0 \ 0 \ 0 \ 2/3 \ 1/3 \ 1/5]$ . Où est l'erreur ?)

► Comment Page et Brin vont-ils obvier aux difficultés rencontrées ? L'idée (plus ou moins explicite) est de changer de point de vue sur ce que signifie la matrice  $\mathbf{H}$ . Considérons par exemple

$$\mathbf{H} = \begin{bmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 0 & 0 & 0 \end{bmatrix}.$$

L'interprétation de départ de la ligne 1 (par exemple) est que la page  $P_1$  « donne » la moitié de son « autorité » aux pages  $P_2$  et  $P_3$ . Considérons alors la matrice  $\mathbf{R}^{(0)T} = [1/3 \ 1/3 \ 1/3]$ . Dans ce cas, où la somme des coefficients est

égale à 1, on peut imaginer un internaute dont la présence sur le Web est gouvernée, au départ (à l'instant 0), par cette distribution (uniforme) de probabilité : à l'instant 0, l'internaute a une chance sur trois d'être sur  $P_1$ , et de même pour la page  $P_2$  ou la page  $P_3$ . S'il est sur  $P_1$ , où sera-t-il à l'instant suivant, l'instant 1 ? On va considérer que la ligne 1 de  $\mathbf{H}$  donne la probabilité que, au coup d'après (au temps 1), il se trouve sur l'une ou l'autre des pages  $P_1, P_2, P_3$ . Cette probabilité est nulle en ce qui concerne la page 1 : on écrira que l'on a

$$P(X_1 = 1 | X_0 = 1) = 0,$$

ce qu'on lit : la probabilité que  $X_1 = 1$ , c'est-à-dire que notre internaute soit (encore) sur  $P_1$  à l'instant 1, sachant qu'il était sur la page  $P_1$  à l'instant 0, est nulle. On aura de même

$$P(X_1 = 2 | X_0 = 1) = 1/2 ; P(X_1 = 3 | X_0 = 1) = 1/2.$$

La probabilité qu'il soit sur  $P_2$  (respectivement, sur  $P_3$ ) à l'instant 1 est donc  $1/2$ .

► De la même façon, on aura :

$$P(X_1 = 1 | X_0 = 2) = 1/2 ; P(X_1 = 2 | X_0 = 2) = 0 ; P(X_1 = 3 | X_0 = 2) = 1/2.$$

Mais que se passe-t-il si, à l'instant 0, l'internaute se trouve sur la page  $P_3$  ? Celle-ci est, dans le Web considéré, un "*dangling node*" (le lecteur traduira par ses propres moyens). C'est là que Page et Brin introduisent l'hypothèse du "*random surfer*", du surfeur au hasard, comme on parle classiquement, en théorie des probabilités, de promeneur au hasard (*random walker*) et de promenade au hasard (*random walk*). L'internaute surfeur au hasard qui se trouve sur la page  $P_3$  et constate que celle-ci n'a pas de liens sortants (*outlinks*) va alors choisir au hasard entre toutes les pages du Web : par hypothèse on posera donc :

$$P(X_1 = 1 | X_0 = 3) = P(X_1 = 2 | X_0 = 3) = 1/3 = P(X_1 = 3 | X_0 = 3) = 1/3.$$

Cela conduit à considérer, non plus la matrice

$$\mathbf{H} = \begin{bmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 0 & 0 & 0 \end{bmatrix}$$

mais la matrice

$$\mathbf{S} = \begin{bmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}.$$

Cette matrice est dite *stochastique* : chacune de ses lignes a, en effet, l'unité pour somme de ses coefficients. Du point de vue du calcul matriciel, on peut l'obtenir de façon générale, pour une matrice  $\mathbf{H}$  à  $n$  lignes et  $n$  colonnes, en introduisant les matrices  $n \times 1$  notée  $\mathbf{e}$  et  $\mathbf{a}$  telle que

$$\mathbf{e}^T = [1 \ 1 \ \dots \ 1 \ \dots \ 1] \text{ et } \mathbf{a}^T = [a_1 \ a_2 \ \dots \ a_i \ \dots \ a_n]$$

où  $a_i = 1$  si la ligne  $i$  de  $\mathbf{H}$  ne comporte que des zéros,  $= 0$  sinon. On a alors :

$$\mathbf{a}(1/n \mathbf{e}^T) = \begin{bmatrix} a_1 \\ a_2 \\ \dots \\ a_i \\ \dots \\ a_n \end{bmatrix} [1/n \ 1/n \ \dots \ 1/n \ \dots \ 1/n] = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1/n & 1/n & 1/n & 1/n & 1/n & 1/n \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

En ajoutant la matrice  $\mathbf{a}(1/n \mathbf{e}^T)$ , à  $n$  lignes et  $n$  colonnes, à la matrice  $\mathbf{H}$ , on obtient donc la matrice  $\mathbf{S}$  :  $\mathbf{S} = \mathbf{H} + \mathbf{a}(1/n \mathbf{e}^T)$ . Le vecteur « binaire »  $\mathbf{a}$  est appelé en anglais *the dangling node vector*.

► Considérons ici la matrice

$$\mathbf{S} = \begin{bmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}.$$

Que se passe-t-il quand on la substitue à  $\mathbf{H}$  dans l'équation  $\mathbf{R}^T = \mathbf{R}^T \mathbf{S}$  donnant le vecteur  $\mathbf{R}$  des rangs ? Ou encore dans la relation d'itération  $\mathbf{R}^{(k+1)T} = \mathbf{R}^{(k)T} \mathbf{S}$  ? On a ici :

$$\mathbf{I} - \mathbf{S} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/3 & 1/3 & 1/3 \end{bmatrix} = \begin{bmatrix} 1 & -1/2 & -1/2 \\ -1/2 & 1 & -1/2 \\ -1/3 & -1/3 & 2/3 \end{bmatrix}.$$

Consultons le logiciel de calcul matriciel Wims déjà utilisé : il indique que le déterminant de  $\mathbf{I} - \mathbf{S}$  est nul, et donc que cette matrice n'est pas inversible. Cela nous laisse donc une chance qu'il existe un vecteur  $\mathbf{R}$  convenable, c'est-à-dire tel que l'on ait  $\mathbf{R}^T = \mathbf{R}^T \mathbf{S}$ .

► Le même logiciel propose de fournir aussi ce qu'on appelle en anglais les *eigenvalues* et les *eigenvectors*, mots étranges, semi-traduction de l'allemand *Eigenwert* et *Eigenvektor*, *Eigen* signifiant « propre », « particulier » : on parle en français de valeurs propres et de vecteurs propres, tout simplement. Une valeur propre d'une matrice  $\mathbf{A}$   $n \times n$  est un nombre  $\lambda$  tel qu'il existe au moins un vecteur  $\mathbf{X}$   $n \times 1$  *non nul* vérifiant l'égalité matricielle

$$\mathbf{A}\mathbf{X} = \lambda\mathbf{X}.$$

Si l'on transpose les deux membres de cette égalité, on obtient l'égalité équivalente

$$\mathbf{X}^T \mathbf{A}^T = \lambda \mathbf{X}^T.$$

Ce qu'il nous faut savoir c'est donc si la matrice  $\mathbf{A}$  telle que  $\mathbf{A}^T = \mathbf{S}$ , soit  $\mathbf{A} = \mathbf{S}^T$ , admet pour valeur propre  $\lambda = 1$ . Puisque

$$\mathbf{S} = \begin{bmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}$$

on a :

$$\mathbf{S}^T = \begin{bmatrix} 0 & 1/2 & 1/3 \\ 1/2 & 0 & 1/3 \\ 1/2 & 1/2 & 1/3 \end{bmatrix}$$

Voici alors le « verdict » du logiciel consulté.

**Matrix calculator**

$$A = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & 0 & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{3} \end{pmatrix}$$

Eigenvalues and eigenvectors of A:

Value	Multiplicity	Vector
-1/2	1	(1,-1,0)
-1/6	1	(1,1,-2)
1	1	(1,1,3/2)

$\mathbf{S}^T$  admet bien la valeur propre 1, avec le vecteur propre  $(1 \ 1 \ 3/2)$ . Posons

$$\mathbf{R}^T = [1 \ 1 \ 3/2].$$

On a bien :  $\mathbf{R}^T \mathbf{S} = [1 \ 1 \ 3/2] \begin{bmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/3 & 1/3 & 1/3 \end{bmatrix} = [1 \ 1 \ 3/2].$

Le vecteur de rangs obtenu convient : moyennant un facteur multiplicatif convenable, égal 1/3,5 en l'espèce, on peut retenir pour vecteur de rangs le vecteur  $\mathbf{R}^T = [1 \ 1 \ 3/2]/3,5 \approx [28,6 \ \% \ 28,6 \ \% \ 42,8 \ \%]$ .

► Qu'en est-il ici du processus itératif défini par la relation  $\mathbf{R}^{(k+1)T} = \mathbf{R}^{(k)T} \mathbf{S}$  ?  
On part de  $\mathbf{R}^{(0)T} = [1/3 \ 1/3 \ 1/3]$  ; voici ce que donne le logiciel utilisé pour  $\mathbf{R}^{(10)T}$ .

**Matrix calculator**

$$A = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

$A * M1^{10} = \begin{pmatrix} \frac{51828151}{181398528} & \frac{51828151}{181398528} & \frac{38871113}{90699264} \end{pmatrix}$

On voit que les deux premiers coefficients sont égaux ; le troisième est égal au premier multiplié par un facteur proche de 1,5, à savoir

$$1,499999990352733208637907997142...$$

Les deux premiers coefficients sont égaux à

$$0,2857142865018176994247715174403...$$

Le troisième vaut, lui,

$$0,4285714269963646011504569651193...$$

Le vecteur  $\mathbf{R}^{(10)T}$  est donc très proche d'un vecteur de la forme

$$[u \ u \ 1,5u] (= u[1 \ 1 \ 3/2]),$$

ce qu'on peut vérifier pour d'autres valeurs de  $k$ . Pour  $k = 20$ , ainsi, on obtient le vecteur  $\mathbf{R}^{(20)T}$  suivant :

$$\left[ \frac{3133850091482551}{10968475320188928} \quad \frac{3133850091482551}{10968475320188928} \quad \frac{2350387568611913}{5484237660094464} \right]$$

Le rapport du 3<sup>e</sup> coefficient au 2<sup>e</sup>, soit

$$\frac{2350387568611913}{5484237660094464} / \frac{3133850091482551}{10968475320188928}$$

est égal à 1,499999999999999840451845045510... Les deux premiers coefficients sont égaux à 0,2857142857142857273100534656726... tandis que le 3<sup>e</sup> vaut

$$0,4285714285714285453798930686547...$$

Ces résultats suggèrent qu'il y aurait convergence vers un vecteur  $[u \ u \ 1,5u]$ , où l'on a :

$$\frac{1}{u} \approx \frac{10968475320188928}{3133850091482551} \approx 3,499999999999999840451845045510...$$

On peut raisonnablement conjecturer que le vecteur limite, qu'on peut noter

$$\mathbf{R}^{(\infty)}, \text{ est donné par : } \mathbf{R}^{(\infty)T} = \frac{1}{3,5} [1 \ 1 \ 1,5] = \left[ \frac{2}{7} \ \frac{2}{7} \ \frac{3}{7} \right].$$

► On aura noté que la somme des coefficients de  $\mathbf{R}^{(\infty)}$  est égale à 1. On aura peut-être observé aussi qu'il semble en être de même pour les vecteurs  $\mathbf{R}^{(10)}$  et  $\mathbf{R}^{(20)}$ . Pour  $\mathbf{R}^{(10)}$ , ainsi, cette somme vaut

$$2 \times \frac{51828151}{181398528} + \frac{38871113}{90699264} = \frac{51828151}{90699264} + \frac{38871113}{90699264} = \frac{90699264}{90699264} = 1.$$

Pour  $\mathbf{R}^{(20)}$ , de même, on a

$$2 \times \frac{3133850091482551}{10968475320188928} + \frac{2350387568611913}{5484237660094464} = \frac{3133850091482551}{5484237660094464} + \frac{2350387568611913}{5484237660094464} = \frac{5484237660094464}{5484237660094464} = 1.$$

En vérité, du fait que l'on est parti d'un vecteur  $\mathbf{R}^{(0)}$  dont la somme des coefficients est 1, et du fait que  $\mathbf{S}$  est une matrice stochastique, il en est de même de  $\mathbf{R}^{(k)}$  pour tout  $k \geq 1$ . Dans le cas le plus général, la démonstration procède par induction sur  $k$ : supposons la propriété acquise pour  $\mathbf{R}^{(k)}$  et montrons alors que cette propriété se transmet à  $\mathbf{R}^{(k+1)}$ . Pour cela, il suffit de

montrer que si un vecteur  $\mathbf{R}^T = [r_1 \ r_2 \ \dots \ r_j \ \dots \ r_n]$  est tel que  $\sum_{j=1}^n r_j = 1$  et si la matrice  $n \times n \ \mathbf{S} = [S_{ij}]$  a chacune de ses lignes, à savoir  $[S_{i1} \ S_{i2} \ \dots \ S_{ij} \ \dots \ S_{in}]$  pour  $i = 1, 2, \dots, n$ , de somme 1, c'est-à-dire si  $\sum_{j=1}^n S_{ij} = 1$  pour tout  $i = 1, 2, \dots,$

$n$ , alors le vecteur  $\mathbf{R}'^T = [r'_1 \ r'_2 \ \dots \ r'_j \ \dots \ r'_n] = \mathbf{R}^T \mathbf{S}$  est tel que  $\sum_{j=1}^n r'_j = 1$ . On a en

effet  $r'_j = \sum_{i=1}^n r_i S_{ij}$  et on aura donc :

$$\sum_{j=1}^n r'_j = \sum_{j=1}^n \left( \sum_{i=1}^n r_i S_{ij} \right) = \sum_{i=1}^n \left( r_i \sum_{j=1}^n S_{ij} \right) = \sum_{i=1}^n r_i \times 1 = \sum_{i=1}^n r_i = 1.$$

► Ce qui précède porte à penser que nous sommes sur une bonne voie. En même temps, nous n'avons pas atteint le but visé, que rappelle les questions énoncées lors de la séance 5 et que l'on reprend ici en remplaçant  $\mathbf{H}$  par  $\mathbf{S}$  (et  $\boldsymbol{\pi}^{(0)T}$  par  $\mathbf{R}^{(0)T}$ ) :

- Will this iterative process continue indefinitely or will it converge?
- Under what circumstances or properties of  $\mathbf{S}$  is it guaranteed to converge?
- Will it converge to something that makes sense in the context of the PageRank problem?
- Will it converge to just one vector or multiple vectors?
- Does the convergence depend on the starting vector  $\mathbf{R}^{(0)T}$  ?
- If it will converge eventually, how long is "eventually"? That is, how many iterations can we expect until convergence?

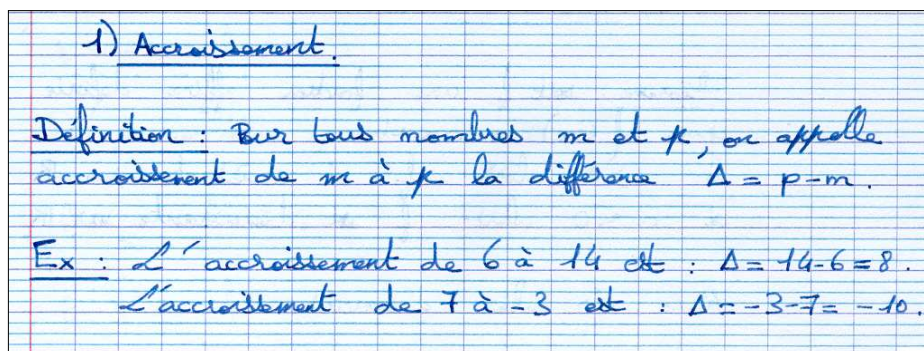
L'invention du surfeur au hasard a permis à Brin et Page de réaliser un premier « ajustement » – l'ajustement de stochasticité (*stochasticity adjustment*). On verra la prochaine fois qu'un second ajustement est nécessaire pour parvenir à un résultat « solide ».

## 2. Quelles mathématiques ? Pour qui ?

a) L'étude de l'algorithme PageRank illustre bien, je crois, ce que je nomme le *souci* des praxéologies *utiles*, souci conçu comme *premier* par rapport aux contraintes usuelles et aux conditions régnantes. Le primat ainsi donné au *critère d'utilité* est de nature à faire apparaître l'arbitraire nécessaire des cursus établis, validés sinon sanctifiés par l'usage, et leur fréquente inadéquation aux besoins praxéologiques les plus légitimes, que ces cursus tendent au demeurant à faire apparaître alors comme illégitimes. Est-il légitime, ainsi, de prétendre savoir quelque chose de la théorie des matrices et de leur « analyse spectrale » (le *spectre* d'une matrice est l'ensemble de ses valeurs propres) quand on s'intéresse aux moteurs de recherche et qu'on n'a pas fait de fortes études de mathématiques ? La réponse ébauchée ci-dessus est : oui, cela l'est ; et c'est par un *raisonnement didactique* – revendiqué de façon parfois arrogante – que nos sociétés sont conduites à nier cette insigne ambition d'une vie meilleure. Le critère d'utilité n'est certes pas facile à manier. Mais en tant qu'instrument de mise en question méthodique de l'état didactique du monde, il est irremplaçable, toujours pertinent, souvent éclairant. J'illustrerai ce thème au cœur de notre travail par deux exemples dont la simplicité égale, à mes yeux, la profondeur.

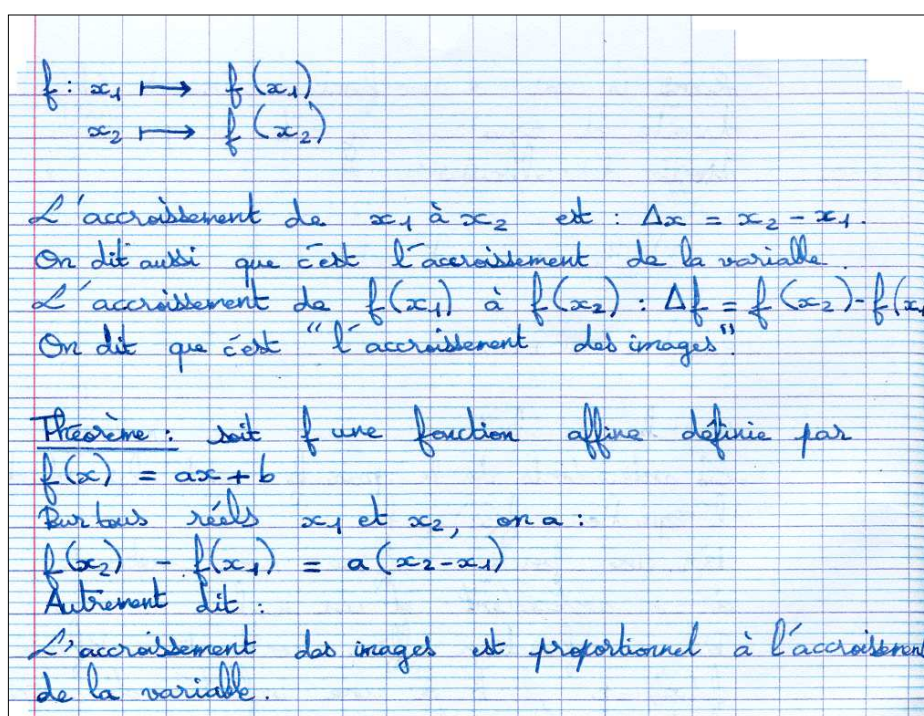
b) Lors de la première partie de la séance 5, l'examen du travail d'un professeur stagiaire avait amené à s'interroger sur un point en apparence mineur relatif aux outils langagiers du travail mathématique.

► De façon typique, un cahier d'élève fait apparaître, dans la classe confiée à ce professeur, ceci (il s'agit d'un extrait du « cours »).



Cette définition identifie « accroissement d'un nombre  $a$  à un nombre  $b$  » et « différence entre  $b$  et  $a$  » : elle institutionnalise ainsi un simple changement de vocabulaire, qui peut être fondé, mais dont la raison d'être n'est à ce stade nullement évidente.

► Le « cours » pris par l'élève se poursuit comme le montre l'extrait ci-après.



On aura observé les variations de notation : alors que la différence  $p - m$  était notée par la lettre  $\Delta$  ( $\Delta = p - m$ ), la différence  $x_2 - x_1$  est maintenant notée  $\Delta x$  ( $\Delta x = x_2 - x_1$ ) : on peut s'interroger sur la raison de cette brusque variation. En même temps,  $x_2 - x_1$  n'est plus désigné comme « accroissement de  $x_1$  à  $x_2$  » mais comme « accroissement de la variable », formulation en



principe incomplète (on pourrait attendre classiquement que l'on parle de « accroissement de la variable entre  $x_1$  et  $x_2$  » ou « de  $x_1$  à  $x_2$  »). De même, ce qui, en conformité avec la définition « générale » posée au départ, est appelé « accroissement de  $f(x_1)$  à  $f(x_2)$  » est aussi désigné, non comme « accroissement de la fonction (entre  $x_1$  et  $x_2$ ) » mais comme « l'accroissement des images ». Là encore, la notation initiale cède la place à une notation nouvelle : cet accroissement est écrit  $\Delta f = f(x_2) - f(x_1)$ . Le théorème qui vient ensuite adopte une formulation « mixte » : « L'accroissement des images est proportionnel à l'accroissement de la variable. »

► L'examen de divers manuels de 2<sup>de</sup> montre que, en la matière, le désordre des mots et des notations règne. Dans le manuel paru chez Hatier en 2004 dans la collection Axiale, on lit ainsi (p. 126) : « Les accroissements de l'image sont proportionnels aux accroissements de la variable. » Ce n'est pas « l'accroissement des images » mais « les accroissements de l'image » : singulier et pluriel ont échangé leur place. Quant à « l'accroissement de la variable » il cède devant *les* « accroissements de la variable ». Pourquoi cela ? Mystère. Un autre ouvrage (Bréal, 2000, p. 118) met tout au singulier, y compris s'agissant de « l'accroissement de l'image » : « Si l'accroissement de l'image est proportionnel à l'accroissement de la variable, alors la fonction  $f$  est affine. » La même « leçon » se retrouve dans le manuel de la collection Fractale publié chez Bordas en 2000, où, à la page 131, on lit ceci : « L'accroissement de l'image par toute fonction affine est proportionnel à l'accroissement de la variable. » En revanche, tel autre manuel (collection Indice, Bordas, 2000, p. 64), dont les auteurs suivent rigoureusement les canons « modernes » et rétablissent ainsi une symétrie ignorée par les manuels cités jusqu'ici, parle d'accroissements non pas « de la variable », mais « des antécédents ». Un passage d'un autre manuel montre mieux encore peut-être la confusion du langage dans le domaine considéré ; le voici (Belin, 2000, p. 90).

Les propositions suivantes sont équivalentes :

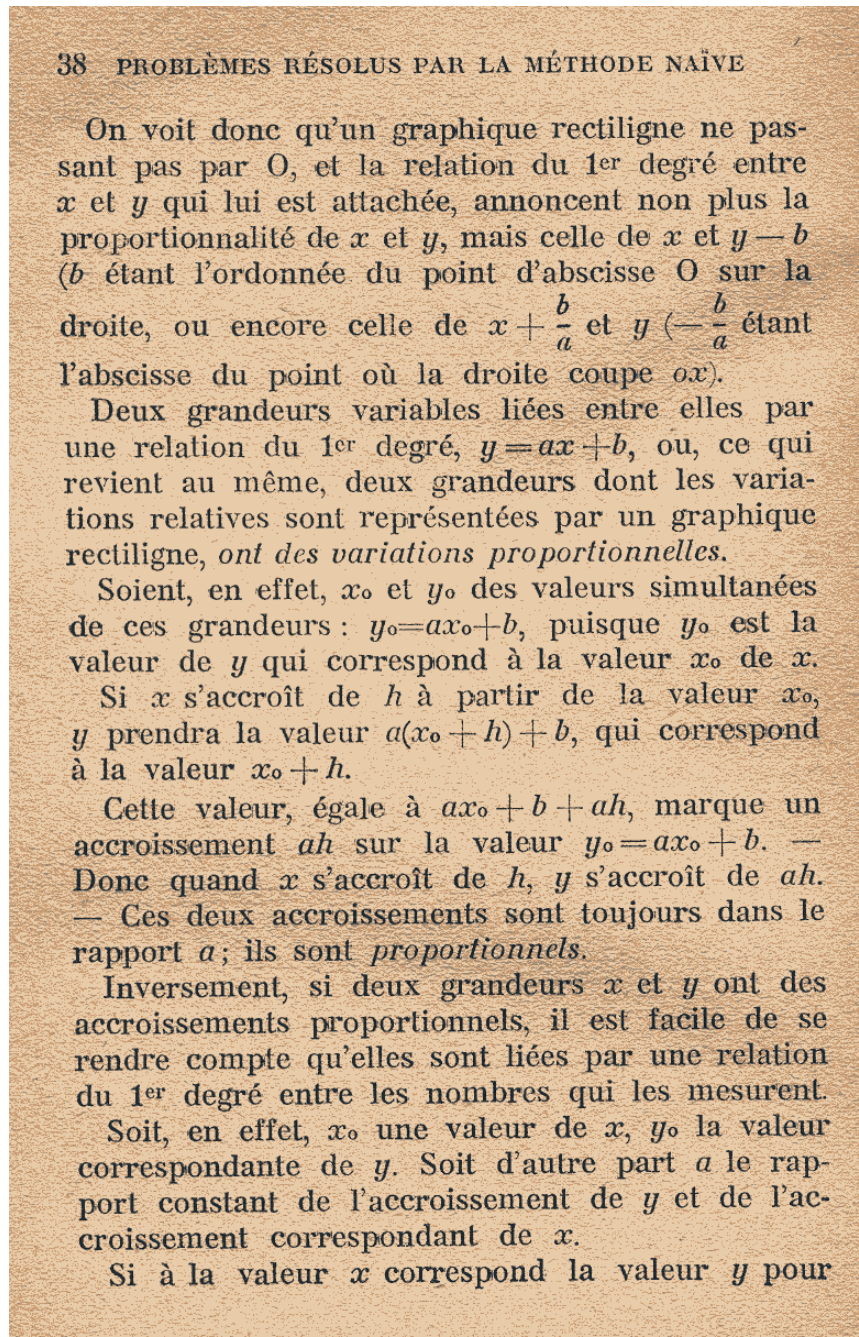
- $f$  est une fonction affine ;
- quels que soient les nombres réels  $a$  et  $b$ , les accroissements  $b - a$  de la variable sont proportionnels aux accroissements  $f(b) - f(a)$  de la fonction ;
- quels que soient les nombres réels distincts  $a$  et  $b$ , le rapport  $\frac{f(b) - f(a)}{b - a}$  est constant.

Pour faire sens, la deuxième proposition devrait être explicitée de façon précise : il existe un nombre réel  $\alpha$  tel que, pour tous  $a, b$ , on a  $f(b) - f(a) = \alpha(b - a)$ .

c) Peut-on expliquer – et peut-être « redresser » – tout cela ? Tel est le *problème de la profession* de professeur de mathématiques que révèlent les observations précédentes. Comment l'aborder ?

► On peut d'abord remonter dans le temps jusqu'à une époque où le vocabulaire « moderne » n'existait pas (et n'était donc pas susceptible de

poser problème). Nous examinerons pour cela deux pages d'un ouvrage intitulé *Les problèmes résolus par la méthode naïve*, qui porte pour sous-titre « Barèmes, graphiques, formules ». Publié en 1929 chez Fernand Nathan, signé de Jules Gal et d'Abel Marijon, l'ouvrage, dont le contenu est résolument élémentaire, est aussi très maîtrisé. Page 38 (ci-après), on verra notamment l'emploi d'un vocabulaire classique, celui des *grandeurs* (variables), notées  $x$  et  $y$  respectivement.



Ici, le système de concepts fait l'économie de la notion moderne de *fonction* : les grandeurs  $x$  et  $y$  sont liées par une *relation* (qui est, il est vrai, une relation *fonctionnelle*). On peut imaginer un système **S** dont l'état dépend du

temps  $t$ , et dont  $x$  et  $y$  sont deux variables d'état, qui dépendent donc du temps : on écrirait aujourd'hui (par exemple)  $x = \varphi(t)$ ,  $y = \psi(t)$  ; on écrivait autrefois :  $x = x(t)$ ,  $y = y(t)$ . Cela noté, on peut « oublier » le temps en ne retenant que la notion de « valeurs simultanées » :  $x_0$  et  $y_0$  sont des valeurs simultanées des (grandeurs) variables  $x$  et  $y$  s'il existe (au moins) un instant  $t_0$  pour lequel on ait :  $x_0 = x(t_0)$  et  $y_0 = y(t_0)$ . On peut dire même que le couple  $(x_0, y_0)$  est un couple de valeurs simultanées s'il existe un état du système  $\mathbf{S}$  pour lequel la (grandeur) variable  $x$  prend la valeur  $x_0$  tandis que la variable  $y$  prend la valeur  $y_0$ .

► La page examinée parle de « variations proportionnelles ». Comment entendre cela ? Supposons un ensemble  $\mathfrak{R}$  de couples de nombres  $(a, b)$ ,  $(a', b')$ ,  $(a'', b'')$ , etc.  $\mathfrak{R}$  est (ou correspond à) une *relation binaire* sur  $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$ . Notons d'abord qu'une telle relation n'est pas nécessairement fonctionnelle : il peut exister des nombre  $a, b, c$ , avec  $c \neq b$ , tels que l'on ait  $(a, b) \in \mathfrak{R}$  et  $(a, c) \in \mathfrak{R}$ . Cela dit, quand dira-t-on que  $\mathfrak{R}$  est une *relation de proportionnalité* (ou *un système proportionnel*, etc.) ? Pour éliminer des cas triviaux, on suppose d'abord que  $\mathfrak{R}$  contient un couple  $(a_0, b_0)$  où  $a_0 b_0 \neq 0$  ; on peut alors adopter la définition suivante : il en est ainsi si, quels que soient  $(a, b) \in \mathfrak{R}$  et  $(a', b') \in \mathfrak{R}$ , on a :  $ab' = a'b$ . Ce critère « régularise » le critère plus intuitif selon lequel, si  $k$  est tel que  $a' = ka$ , alors on a  $b' = kb$ , qui pose problème lorsque  $a = 0$  et  $a' \neq 0$ .) On peut alors observer ceci : si  $\mathfrak{R}$  est une relation proportionnelle, alors  $\mathfrak{R}$  est une relation *fonctionnelle* : si, en effet,  $a = a' \neq 0$ , alors l'égalité  $ab' = a'b$  entraîne  $b = b'$  ; en d'autres termes, on ne peut avoir dans  $\mathfrak{R}$  de couples  $(a, b)$  et  $(a, b')$  tels que  $b \neq b'$ . (On laissera à la sagacité du lecteur l'examen du cas où  $a = a' = 0$ .) Surtout, on a le résultat suivant : si  $\mathfrak{R}$  est une relation de proportionnalité, il existe un nombre  $\alpha$  tel que, quel que soit le couple  $(a, b) \in \mathfrak{R}$ , on a :  $b = \alpha a$ . Pour le voir, il suffit de poser  $\alpha = b_0/a_0$  ; l'égalité  $a_0 b = ab_0$  entraîne alors que l'on a :  $b = (ab_0)/a_0 = (b_0/a_0)a = \alpha a$ .

► Dans le cas qui nous occupe, les couples de nombres sont des « variations » de  $x$  et de  $y$  respectivement. Si  $(x_1, y_1)$  et  $(x_2, y_2)$  sont des couples de « valeurs simultanées » des variables  $x$  et  $y$ , un couple de variations est par exemple le couple  $(x_2 - x_1, y_2 - y_1)$ . Notons ici  $\mathfrak{R}(x, y)$  l'ensemble de ces couples ; le texte examiné affirme alors que

1) s'il existe une relation du 1<sup>er</sup> degré entre les variables  $x$  et  $y$ , alors la relation  $\mathfrak{R}(x, y)$  est de proportionnalité : il existe  $\alpha$  tel que, quels que soit les couples  $(x_1, y_1)$  et  $(x_2, y_2)$  de « valeurs simultanées » des variables  $x$  et  $y$ , on a

$$y_2 - y_1 = \alpha (x_2 - x_1).$$

2) Réciproquement, si la relation binaire  $\mathfrak{R}(x, y)$  est proportionnelle, alors les variables  $x$  et  $y$  sont liées par une relation du 1<sup>er</sup> degré.

► La réciproque précédente peut être renforcée en affaiblissant l'hypothèse sur laquelle elle s'appuie : au lieu de l'ensemble

$$\mathfrak{R}(x, y) = \{ (x_2 - x_1, y_2 - y_1) / \text{pour tous les couples } (x_1, y_1) \text{ et } (x_2, y_2) \text{ de valeurs simultanées de } x \text{ et } y \}$$

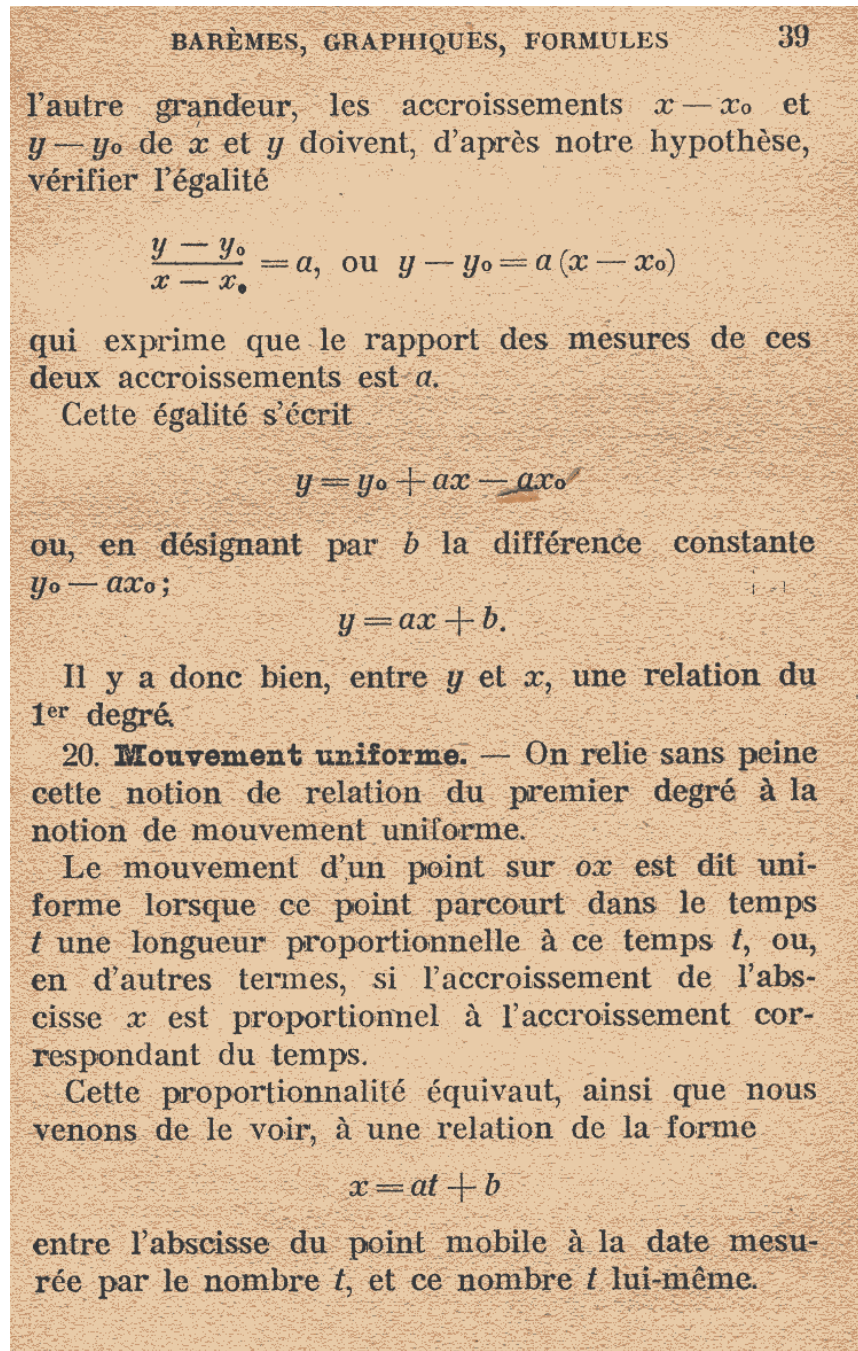


il suffit de supposer que,  $(x_0, y_0)$  étant un couple de valeurs simultanées de  $x$  et  $y$ , la relation binaire

$$\mathfrak{R}(x, y; x_0, y_0) = \{ (x_1 - x_0, y_1 - y_0) / \text{pour tous les couples } (x_1, y_1) \text{ de valeurs simultanées de } x \text{ et } y \}$$

est de proportionnalité. Il existe en effet  $\alpha$  tel que l'on ait  $y_1 - y_0 = \alpha (x_1 - x_0)$  pour tout couple  $(x_1, y_1)$  de valeurs simultanées, en sorte que  $y_1 = \alpha x_1 + y_0 - \alpha x_0$ , ce qui s'écrit encore :  $y = \alpha x + y_0 - \alpha x_0$ .

► Passons maintenant à la suite de la page 38 examinée.



Ce qui est surtout remarquable, ici, c'est le contenu de la section *Mouvement uniforme*, qui montre clairement (« Cette proportionnalité équivaut, ainsi que nous venons de le voir, à une relation du type... ») le rôle assigné à la partie « réciproque » du théorème de caractérisation des fonctions affines dans le « texte du savoir mathématique » classique – un rôle apparemment totalement oublié lors de la résurgence de cette réciproque dans le programme de 2<sup>de</sup> entré en vigueur à la rentrée 2000 : lorsque des variables  $x$  et  $y$  ont des « variations proportionnelles », alors elles sont liées par une relation du premier degré. Un robinet déverse de l'eau dans une cuve ; le débit de ce robinet est *constant* : il y a donc « variations proportionnelles » entre durée  $t$  du versement et volume  $V$  contenu dans la cuve. Il existe par conséquent des nombres  $a$  et  $b$  tels que, pour tout  $t$ , on ait  $V = at + b$ . Les exemples pourraient évidemment être multipliés.

d) Que faire de la notation  $\Delta$  ? C'est là que la notion de *fonction* va entrer en scène.

► Voici d'abord un petit rappel que Frédéric Pham, commentant la notion de différentielle dans son livre *Géométrie et calcul différentiel sur les variétés* (InterEditions, Paris, 1992, p. 23), est amené à formuler.

### 0.3 Sur les notations en mathématiques

Le problème de la notation différentielle n'est qu'un aspect d'un problème beaucoup plus vaste de notations en mathématiques. Très schématiquement, on peut opposer le système de notations « classique », qui « déclare » les *variables* en leur attribuant des noms  $x$ ,  $y$ , ... (et n'éprouve pas toujours le besoin d'attribuer des noms aux relations entre ces variables), et le système de notations « moderne » qui ne déclare que les *relations* (par exemple les applications), prenant soin de noter différemment même des applications trivialement déduites l'une de l'autre (par exemple une application et sa restriction), et de donner des noms même à des applications « triviales » comme l'inclusion d'un ensemble dans un autre.

Ce deuxième système de notations est devenu le « système officiel » dans la remise en ordre des mathématiques qui caractérise le milieu de ce siècle. Il a certainement beaucoup aidé à dissiper certaines confusions du discours mathématique classique. Mais conçu avant tout pour les grandes synthèses abstraites, il est souvent lourd à utiliser dans la pratique (et complètement inadapté, dans la plupart des cas, à la façon de penser des physiciens, pour ne citer qu'eux). Dans leur pratique quotidienne les mathématiciens utilisent donc tour à tour l'un ou l'autre système de notations, ou un mélange des deux. Mais au lieu de chercher à enseigner aux étudiants les mathématiques telles qu'elles se pratiquent, on continue trop souvent à propager le mythe d'une notation mathématique « parfaite » et unique, indépendante du contexte, permettant d'éviter « automatiquement » de faire des erreurs de raisonnement – mythe dont le succès est d'autant mieux assuré qu'on s'arrange pour ne présenter aux étudiants que des abstractions isolées de tout contexte, où effectivement le mythe fonctionne bien. Moyennant quoi les étudiants confondent l'apprentissage, de toutes façons difficile, de la rigueur, avec la soumission aveugle à une « langue de bois » qui paralyse toute pensée.

► On sait que, dans le paradigme classique, on écrivait une fonction  $f$  sous la forme  $y = f(x)$ . Le symbole  $f$  ne faisait ici que dénoter la procédure permettant de passer d'une valeur de la *variable*  $x$  supposée à la valeur

« simultanée » de la *variable y* supposée. (En de nombreux domaines d'application,  $x$  est appelée *variable indépendante*, tandis que  $y$  est la *variable dépendante*.) Par contraste, dans le paradigme moderne, on « évide » la référence à des variables  $x$  et  $y$  qui seraient définies *par ailleurs*. Le symbole  $f$  désigne le « protocole » permettant d'associer une valeur numérique à une autre, ces valeurs n'étant alors plus valeurs *de rien* : ce sont des valeurs « abstraites », des nombres purs, qui ne « mesurent » rien. Comme il faut bien noter le nombre  $x$  auquel  $f$  associe le nombre  $f(x)$ , la lettre  $x$ , et son nom ancien, « variable », ont été conservés ; mais  $y$  a été banni, au profit du schéma notationnel « bien connu »  $f : x \mapsto f(x) = \dots$ . Là où l'on parlait autrefois « de la fonction  $y = 2x^3 - 5$  » par exemple, on s'est mis à parler de « la fonction  $f$  qui, à  $x$ , associe  $2x^3 - 5$  », de « la fonction  $f$ , " $x$  donne  $2x^3 - 5$ " », etc.

► Voyons alors quel usage peut être fait du symbole  $\Delta$ . Nous suivrons ici un ouvrage un peu ancien, intitulé *Difference Equations*, dû à Samuel Goldberg et publié 1958 chez John Wiley & Sons, Inc. (New York). Le passage suivant est extrait des pages 13-14.

... if  $y$  denotes a population function, with  $y(x)$  being the size of the population in census year  $x$ , we may define a new function whose value at  $x$  is the difference between the population size in years  $x$  and  $x + 10$ . Since this new function is related to  $y$ , we call it  $\Delta y$ , so that [with  $\Delta y(x)$  as the value of  $\Delta y$  at  $x$ ]

$$\Delta y(x) = y(x + 10) - y(x).$$

$\Delta y(x)$  is the change in population size in the two consecutive census years  $x$  and  $x + 10$ . If the population has increased, then  $\Delta y(x) > 0$ ; if the population has declined, then  $\Delta y(x) < 0$ ; if the population size is unchanged, then  $\Delta y(x) = 0$ .

Similarly, if  $C$  is the function whose value at  $t$  is denoted by  $C_t$ , then we may define another function  $\Delta C$  by specifying that its value at  $t$  is the difference  $(C_{t+2} - C_t)$ . If the value of  $\Delta C$  at  $t$  is denoted by  $\Delta C_t$ , then

$$\Delta C_t = C_{t+2} - C_t$$

and we may choose to interpret this as the difference between consumer expenditures in the 2-year interval between periods  $t$  and  $t + 2$ .

These considerations are formalized in the following definition:

DEFINITION 1.2. Let a function  $y$  be given and let  $h$  be any constant for which  $x + h$  is in the domain of  $y$  whenever  $x$  is. Then  $\Delta y$ , the first difference of  $y$ , is that function whose value at  $x$ , denoted by  $\Delta y(x)$  (or  $\Delta y_x$ ), is given by

$$(1.3) \quad \Delta y(x) = y(x + h) - y(x).$$

The symbol  $\Delta$  denotes the *difference operator*, indicating that the function  $y$  is to be operated upon (or transformed) to yield the new function  $\Delta y$ . The number  $h$  is called the *difference interval*.

► Plusieurs aspects méritent d'être soulignés. D'abord le symbole  $\Delta$  désigne ici un certain *opérateur*, qui transforme une *fonction* en une autre *fonction*, ce qu'on pourrait noter en style moderne  $\Delta : f \mapsto \Delta f$ . La notation  $\Delta y$  (ou, aujourd'hui,  $\Delta f$ ) désigne donc une *fonction*. Ce qu'il faut noter ensuite, c'est

que, ici, la fonction  $\Delta y$  est une fonction d'une variable  $x$ , et cela grâce à un subterfuge consistant à introduire un paramètre  $h$ , le *difference interval*. L'auteur que nous suivons ajoute à ce propos cette remarque, qui fera avancer notre travail.

We shall understand, unless it is specifically indicated otherwise, that we are always using the same difference interval  $h$ . Because of this agreement, we shall not explicitly indicate the dependence of  $\Delta y$  on  $h$ . (Some authors write  $\Delta_h y$  in order to avoid all possibility of misunderstanding.) For the simple function given by  $y(x) = x$ , we find, using (1.3),

$$\Delta y(x) = \Delta x = (x + h) - x$$

or

$$\Delta x = h.$$

For this reason, the difference interval is often denoted by  $\Delta x$ .

► On peut tirer de ce qui précède un certain nombre de conclusions. Soit  $f$  une fonction d'une variable  $x$  ; alors  $\Delta f$  est, en principe, une fonction de deux variables  $x_1$  et  $x_2$  :  $\Delta f : (x_1, x_2) \mapsto \Delta f(x_1, x_2) = f(x_2) - f(x_1)$ . Selon la tradition, «  $f(b) - f(a)$  s'appelle *accroissement* de  $f$  entre  $a$  et  $b$  », notent par exemple Charles Pisot et Marc Zamanski dans leur ouvrage *Mathématiques générales* (Dunod, Paris, 1963, p. 291), dans le cadre d'un commentaire relatif au « théorème des accroissements finis ». (Ils ajoutent au reste que cette appellation suppose que l'on ait  $a < b$ .) L'accroissement de  $f$  entre  $x_1$  et  $x_2$  est donc la valeur sur le couple  $(x_1, x_2)$  de la fonction  $\Delta f$ . Lorsque  $f$  est la fonction *identité*, notée  $\text{Id}$ , telle que  $\text{Id}(x) = x$  pour tout  $x$ , on a :

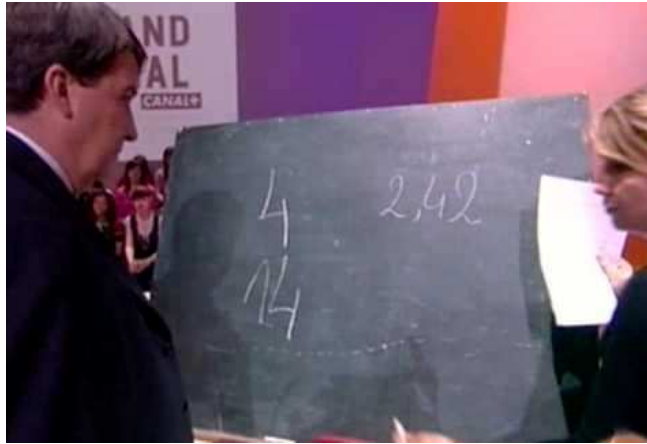
$$\Delta \text{Id}(x_1, x_2) = \text{Id}(x_2) - \text{Id}(x_1) = x_2 - x_1.$$

Comment justifier la notation  $\Delta x$  ? Si l'on parle de fonction  $y = f(x)$ , on peut être amené à écrire, comme le fait S. Goldberg plus haut,  $\Delta y$  en lieu et place de  $\Delta f$  ; comme  $x = \text{Id}(x)$ , cela conduit à écrire  $\Delta x$  en lieu et place de  $\Delta \text{Id}$ . Le fait d'écrire  $\Delta f$  plutôt que  $\Delta y$  et – inversement –  $\Delta x$  plutôt que  $\Delta \text{Id}$  n'est pas en soi « condamnable » ; mais il faut se souvenir alors que l'on n'a pas, sauf par abus d'écriture,  $\Delta f = f(x_2) - f(x_1)$ , mais bien :  $\Delta f(x_1, x_2) = f(x_2) - f(x_1)$  ; et que, de même, on a  $\Delta x(x_1, x_2) = x_2 - x_1$ , plutôt que  $\Delta x = x_2 - x_1$ , écriture dans laquelle la lettre  $x$  figurant dans  $\Delta x$  désigne la fonction identité, non la variable  $x$  elle-même. Tout cela noté (et compris), on pourra, pour alléger l'écriture sans pour autant oublier de quoi il retourne en tout cela, écrire *par exemple* ceci, que je ne crois pas hors de portée d'élèves de seconde :

$$\Delta f_{x_1}^{x_2} = f(x_2) - f(x_1) \text{ et } \Delta x_{x_1}^{x_2} = [x(x_2) - x(x_1)] = x_2 - x_1.$$

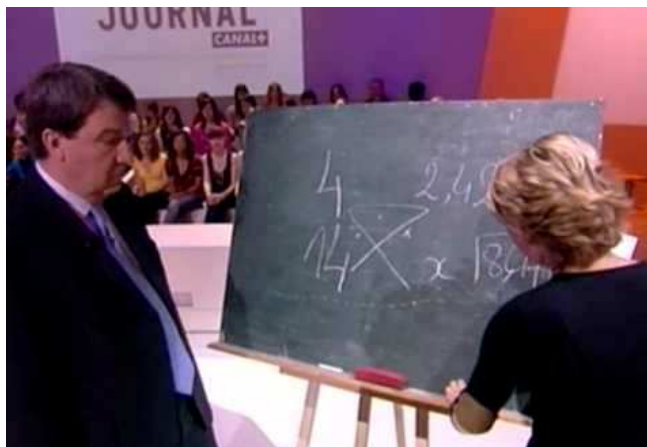
e) Je voudrais ajouter à l'exemple précédent un exemple qui ne concerne pas uniquement la profession de professeur de mathématiques, mais bien tous les citoyens. L'actuel ministre de l'Éducation nationale, Xavier Darcos, s'est illustré récemment (le 3 avril 2008) dans l'émission *Le Grand Journal* de la chaîne Canal +, par son incapacité à résoudre un problème simple dit « de règle de trois ».

► La photo ci-après montre X. Darcos face au tableau où la chroniqueuse Ariane Massenet a écrit les données du problème en question : « Sachant que 4 stylos valent 2,42 €, combien valent 14 stylos ? » D'emblée, Xavier Darcos lui lance : « Je ne sais pas le faire du tout. Comment faites-vous ? Montrez-moi ! » (Voir la vidéo à l'adresse <http://www.canalplus.fr/index.php?pid=2438>.)



► Ayant vu le ministre en difficulté sur une question de grammaire (conjuguer le verbe *aimer* au passé antérieur, à la première personne du pluriel), et pour excuser à l'avance la piètre prestation attendue du ministre sur le problème des stylos, le chroniqueur Jean-Michel Apathie lance : « Il est agrégé de lettres, je crois ; il n'est pas mathématicien... » Il faudrait donc être « mathématicien » pour calculer le prix de 14 stylos connaissant le prix de 4 stylos ! Et sans doute est-ce presque strictement vrai : on mesure à cela combien l'équipement praxéologique que prétend diffuser l'école, à cet égard, demeure à l'intérieur des jeux de l'école et ne mord guère à l'extérieur, dans les labours de la cité !

► Dans le cas d'espèce, on voit mal comment la technique des « produits en croix », lourdement scolaire, mise en œuvre de manière stylisée par Ariane Massenet pourrait trouver sa place dans la cité. Rappelons-la : on multiplie par 14 les 2,42 € et on divise par 4 ; on trouve 8,47 €, comme le montrent les inscriptions figurant au tableau (voir ci-dessous) et comme le précise le dialogue ci-après.





- On va l'appeler  $x$ , d'accord ? On va multiplier 14 par 2 euros 42...
- Certes.
- C'est le signe de la règle de trois : là, on multiplie, et là on divise... par 4. Vous comprenez ?

► En fait, Ariane Massenet indique tout aussitôt que l'on peut procéder de façon beaucoup plus simple, par ce qui n'est autre que la technique *de réduction à l'unité* : on divise 2,42 € par 4 et on multiplie par 14. Or sans doute personne – ni elle, ni ses « conseillers techniques », qui sont peut-être des professeurs de mathématiques – n'a véritablement mis en œuvre cette technique ici : car si l'on divise 2,42 par 4, on trouve la moitié de 1,21, à savoir 0,605, nombre qui ne saurait être un prix unitaire en euros ! La non-diffusion d'une technique adéquate va ici de pair, presque exemplairement, avec une indifférence ambiguë aux conditions ordinaires dans lesquelles pourrait être mise en œuvre une technique mieux adaptée aux contraintes de la cité. La fragilité, la volatilité, l'intransportabilité des techniques scolairement enseignées est ici évidente.

► Dans l'ouvrage de Jules Gal et Abel Marijon cité plus haut, on lit également ceci (p. 26 ; ce sont les auteurs qui soulignent) : lorsqu'il y a proportionnalité entre deux grandeurs [= variables], « si l'on multiplie l'une des grandeurs par un nombre quelconque, *entier ou fractionnaire*, la grandeur correspondante est multipliée par le même nombre ». Par exemple, si l'on passe de 4 stylos à 14 stylos, le nombre de stylos est multiplié par  $\frac{14}{4}$  ; par suite, le prix des stylos, soit 2,42 €, est multiplié par  $\frac{14}{4}$  : 14 stylos valent donc  $\frac{14}{4} \times 2,42$  €. En pratique, une technique simple et robuste pourrait donc consister à dire que « si 4 stylos coûtent 2,42 €, 14 stylos coûteront  $\frac{14}{4}$  fois plus, soit  $\frac{14}{4} \times 2,42$  € », avant de confier ce dernier calcul à une calculatrice. Or on doit constater que cette technique – qui mobilise, certes, le « difficile » concept de nombre *fractionnaire* de « fois » – n'a pas reçu l'attention généreuse que l'on doit de façon générale aux moyens de satisfaire effectivement les besoins praxéologiques des gens...

## **SUR LA TAD**

### **1. Questions de style**

a) Je voudrais aujourd'hui me hasarder à proposer l'ébauche du système de références documentaires qui serait partie prenante du « style TAD ». Ce système s'appuie fortement sur celui de l'APA, avec des écarts. La source principale sur le style de l'APA est le *Publication Manual* dans sa 5<sup>e</sup> édition de 2001 : celui-ci sera désigné par le sigle PM dans la suite.

b) Je considère d'abord la question de ce que je nommerai ici les *appels de références*. Dans son chapitre 3, *APA Editorial Style*, le PM lui consacre une section intitulée *Reference Citations in Text*, qui va de la page 207 à la page 214.

c) Tout de suite se présente un point sur lequel le style TAD se distingue du style APA : on ne s'y réfère jamais (sauf dans les appels entre parenthèses) au seul *nom* d'un auteur, comme il en va dans le passage suivant (p. 207) :

Walker (2000) compared reaction times

On adopte en effet la règle suivante : lors du *premier* appel d'une référence dans l'*article* ou le *chapitre* d'ouvrage considéré, on indique le *prénom* des auteurs dont le nom est mentionné. On aura ainsi par exemple :

Jimmy Walker (2000) compared reaction times

ou plutôt

Jimmy Walker (2000) a comparé les temps de réaction

► Si le nom de Jimmy Walker a déjà été mentionné, on n'indiquera ensuite que l'initiale du prénom (suivie d'un point et d'une espace insécable) :

J. Walker (2000) a également étudié la composition des extraits

Si l'on doit citer un autre document signé du même auteur supposé déjà mentionné, on écrira encore, par exemple :

J. Walker (2004) a pu confirmer les résultats obtenus par Robert Douglas (2002)

(Dans le cas précédent, on suppose que le nom de Robert Douglas apparaît pour la première fois.)

► L'initiale du prénom disparaît lorsque l'appel de référence figure entre parenthèses, ainsi qu'il en va dans ce deuxième exemple du PM :

In a recent study of reaction times (Walker, 2000)

ou plutôt

Dans une étude récente des temps de réaction, (Walker, 2000)

Mais cela n'est possible que si l'auteur a déjà été appelé. S'il n'en est pas ainsi, on modifiera la phrase de façon adéquate, par exemple ainsi :

Dans une étude récente des temps de réaction, Jimmy Walker (2000)

On n'écrira donc jamais

In 2000 Walker compared reaction times

ou plutôt

En 2000 Walker a comparé les temps de réaction

On aura soit (1<sup>re</sup> occurrence)

En 2000 Jimmy Walker a comparé les temps de réaction

ou (selon le cas)

En 2000, J. Walker a comparé les temps de réaction

► Le PM précise encore que la date peut être omise dans un *paragraphe* dans lequel le document a déjà été cité, comme ici :

In a recent study of reaction times, Walker (2000) described the method.... Walker also found

On écrira

Dans une étude récente des temps de réaction, J. Walker (2000) a décrit la méthode... J. Walker a également établi

ou (s'il s'agit de la première occurrence du nom)

Dans une étude récente des temps de réaction, Jimmy Walker (2000) a décrit la méthode... J. Walker a également établi

d) Le PM examine ensuite le cas *One Work by Multiple Authors*.

► Dans le cas de *deux* auteurs, il convient de citer les deux noms à chaque occurrence dans le texte (en appliquant la règle précédente propre au style TAD) : au lieu de

Arson, and Rock (1994) found

on écrira

Joe Arson et Robert Rock (1994) ont établi

si ces deux noms apparaissent l'un et l'autre pour la première fois ou

J. Arson et Robert Rock (1994) ont établi

si le premier auteur a déjà été mentionné dans le texte tandis que le second est mentionné pour la première fois, enfin

J. Arson et R. Rock (1994) ont établi

si l'un et l'autre auteurs ont déjà été mentionnés dans le texte. Bien entendu on peut écrire aussi

Il est établi (Arson & Rock, 1994) que

► Dans le cas de *trois*, *quatre* ou *cinq* auteurs, tous les noms sont cités lors de la première occurrence ; on aura ainsi

Joe Arson, Helena Murray et Robert Rock (1999) ont établi

ou selon le cas

J. Arson, H. Murray et Robert Rock (1999) ont établi

etc. Lors des mentions suivantes *dans le même paragraphe*, on ne garde que le premier auteur :

J. Arson et al. (1999) ont également établi

(Le « et al. » est en romain.) Dans le cas où cette opération donne lieu à la même forme abrégée, comme il en va avec

J. Arson, H. Murray et Robert Rock (1999) ont établi

et

J. Arson, R. Rock et Paul Cocker (1999) ont décrit avec précision  
qui donnent l'une et l'autre

J. Arson et al. (1999)...

dans chaque appel on mentionnera les auteurs successifs jusqu'à pouvoir  
distinguer les deux appels l'un de l'autre :

J. Arson, H. Murray et al. (1999) ont établi

et

J. Arson, L. Bradley et al. (1999) ont décrit avec précision

► Quand un document a six auteurs ou plus, on ignore les auteurs à partir  
du sixième en écrivant « et al. » à la suite du cinquième nom, sauf lorsque  
une confusion est possible (comme ci-dessus). On aura donc dans le cas  
général, pour la première citation dans un paragraphe :

Joe Arson, Helena Murray, Robert Rock, Denis Thorn, Jean Waltzer et al. (2002) ont  
établi

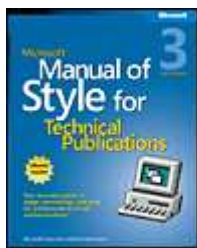
► On notera aussi

Il est établi (Arson, Murray & Rock, 1999) que

et encore

Il est établi (Arson, Murray, Rock, Thorn, Waltzer et al., 2002) que

e) D'autres cas restent à examiner : nous y viendrons la prochaine fois. Mais  
je voudrais ici introduire – rapidement – une hypothèse dont la fragilité  
n'enlève rien à son pouvoir de suggestion, qui en justifie à mes yeux la  
formulation.



► Je pars pour cela d'une observation très simple. Dans  
la floraison de guides de style en langue anglaise, on  
trouve notamment un *Microsoft Manual of Style for  
Technical Publications*, dont la 3<sup>e</sup> édition comporte 432  
pages. Il n'est pas inintéressant d'examiner ce passage de  
la notice correspondante, en ligne sur le site de Microsoft  
(<http://www.microsoft.com/mspress/books/6074.aspx>).

Get clear, concise guidance to help you:

- Use technical terms correctly and consistently—including do's, don'ts, and alternatives for usage.
- Employ the appropriate tone and voice for your audience.
- Produce written and visual content suitable for a worldwide audience.
- Apply best practices for writing and tagging Web content.
- Write better documentation—from dialog boxes and error messages to Web pages and software code.
- Know the standards for creating accessible communications.
- Optimize your indexes, cross-references, and keyword lists.
- Get fast answers on spelling, grammar, and punctuation.

► Ce que je veux souligner par cette indication, c'est l'extrême attention accordée, dans les pays concernés, à la langue anglaise, et notamment à la langue *écrite*, élargie à des usages qui, hier, n'existaient pas, et cela dans tous ses détails : grammaire, orthographe, ponctuation apparaissent ici comme *des savoirs professionnels*, par contraste avec la vision qui prévaut encore en France, celle de savoirs puérocentrés (ou, pour l'adulte, de loisir studieux), ou bien de savoirs spécialisés (du lexicographe, du grammairien, du « linguiste », etc.).

► L'inattention à la langue (française) *écrite*, qui semble prévaloir aujourd'hui chez certaines élites de langue française, me paraît être le symptôme ubiquitaire d'un phénomène de longue durée sans doute, mais non moins ravageur : je veux parler de la lente, incertaine, d'abord inapparente transformation d'une langue en *patois*, un « patois » qu'on ne sait plus *écrire*, ou plutôt qu'on écrit selon sa fantaisie. L'histoire de l'occitan fournit de cela, à l'évidence, un douloureux exemple ; elle nous apprend aussi que tout ne se joue pas, certes, dans le rapport à la langue, et que celui-ci découle de contraintes plus fortes, de longue durée, consubstantielles à une domination politique, économique, culturelle et, en fin de compte, linguistique. *Mais on peut ou non accepter ces contraintes*. On peut ou non accepter d'entériner personnellement une pratique *patoise* de la langue française.

► Bien entendu, une autre position est possible : celle d'une forte attention à la langue, dégagée de tout passéisme (mais historiquement informée), étrangère à tout esthétisme ambigu (mais soucieuse de façonner un bon outil de langage). Ce *souci de la langue* participe à mes yeux d'un engagement « citoyen », au service de la vie bonne de tous ceux qui ont à s'exprimer en français ou choisissent de le faire. Qu'est-ce, en effet, qu'une langue à propos de laquelle les locuteurs natifs, et même les « experts » supposés en la matière, consultés, ne savent répondre aux questions de style que : « On ne sait pas... Faites comme bon vous semble ! » ? Une telle langue, serait un *instrument défectueux*, auquel on préférera bientôt un instrument réglé, précis, sans cesse « travaillé » en vue d'un usage mieux fondé, davantage assuré. *L'engagement « anti-patoisant »* à l'endroit du français suppose, à l'inverse, de prendre d'abord conscience de l'exigence d'un engagement *professionnel* dans des savoirs de la langue – orale et écrite – qui, jusqu'ici, sont encore regardés comme un horizon lointain, dont nous avons appris à détourner nos regards. Un apprentissage collectif inédit est ainsi de mise.

## **2. E-learning et TAD**

a) Je me suis engagé à m'exprimer, dans le cadre d'un numéro à paraître d'une revue italienne, le *Journal of e-Learning and Knowledge Society* (voir <http://www.je-lks.it/>), à propos du *e-learning*, saisi avec les outils de la TAD. Je m'arrête ici sur ce terme ainsi que sur quelques autres qui lui sont liés.

► Quand on saisit **e-learning** dans la fenêtre de recherche du site FranceTerme (<http://franceterme.culture.fr/FranceTerme/recherche.html>), site « consacré aux termes recommandés au *Journal officiel de la République française* », on trouve l'expression « formation en ligne », qui, en anglais, se rendrait, selon cette source, par *e-learning*, *e-training* ou *on-line training* (JO du 14/05/2005).

► L'article « Apprentissage en ligne » de Wikipédia fournit quelques repères non dépourvus d'utilité. J'en reproduis un premier passage – qui illustre, en passant, une certaine pratique patoise de la langue française écrite.

#### **Définition**

La définition de l'apprentissage en ligne (*e-learning*) donnée par l'Union Européenne est : « l'e-learning est l'utilisation des nouvelles technologies multimédias de l'Internet pour améliorer la qualité de l'apprentissage en facilitant d'une part l'accès à des ressources et à des services, d'autre part les échanges et la collaboration à distance ».

En anglais, le terme E-learning, employé par le monde économique, résulte d'une volonté d'unifier des termes tels que : « Open and Distance Learning » (ODL) pour qualifier sa dimension ouverte et qui vient du monde de la formation à distance, « Computer-Mediated Communication » (CMC) pour traduire les technologies de communication (Mails, Forum, Groupware) appliquées à la formation « Web-Based Training » (WBT) pour traduire la technologie dominante sur Internet pour la formation, « Distributed Learning » qui traduit plus une approche pédagogique de type constructiviste et fondée sur la Cognition Distribuée (Grabinger *et al.*, 2001).

► Peut-on dégager, sinon une unité, du moins une ébauche de classification de la diversité des pratiques évoquées dans les textes sur le e-learning ? Voici déjà un petit élément de réponse, sous la forme d'un autre passage de l'article déjà cité : on y notera une interpolation facétieuse, du genre de celles qui font dire aux esprits rassis que Wikipédia « n'est pas fiable » et autres billevesées.

#### **Typologie de l'apprentissage en ligne**

Différentes typologies cohabitent avec Mireille et Robert dans l'apprentissage en ligne. On peut définir une formation en ligne selon deux axes :

\* L'axe de l'espace formateur - apprenant. Il peut exister une grande proximité, lorsque le professeur donne un cours devant ses étudiants, on parle de formation en présentiel. Lorsque le formateur n'est pas en contact avec l'apprenant, on parlera de formation à distance ou de formation distancielle.

\* L'axe du temps de la communication. Si la communication est directe, immédiate, on parle d'outil synchrone. C'est le cas du professeur face à ses étudiants mais aussi lors d'une séance de « chat » ou une visio-conférence. Lorsqu'un délai existe entre une question et sa réponse on parlera d'outil asynchrone. Le forum, le courrier électronique sont des outils de communication asynchrone.

► Jusqu'ici, notons-le, les différences entre ce tableau et celui qu'on aurait pu brosser *avant* l'Internet ne sont pas frappantes. L'article que je suis ici précise ensuite le cas où il y a « 100% apprentissage en ligne ».

Dans cette typologie de formation, l'apprenant va suivre sa session d'apprentissage en ligne entièrement à distance. Il n'interagira pas directement en vis-à-vis avec le ou les formateurs. Il pourra néanmoins bénéficier de [tutorat à distance](#) via des outils de type webconférencing, forum, chat, téléphone, email.

On voit apparaître un personnage un peu énigmatique (nous y reviendrons), le *tuteur à distance*, distinct du formateur, mais qui, en tant que tel, n'est pas davantage rencontré physiquement par l'étudiant en ligne (même s'il peut communiquer avec lui « en mode synchrone » : chat, téléphone, etc.).

► En contraste avec l'apprentissage « 100 % en ligne », il y a l'apprentissage *mixte*, présenté dans les termes suivants.

Le mode d'apprentissage mixte (appelé aussi Blended Learning en anglais) désigne l'utilisation conjointe du eLearning et du mode classique d'apprentissage appelé souvent « présentiel ». En général, l'apprenant va ainsi alterner entre des sessions à distance en ligne et des sessions en face-à-face avec le ou les formateurs. Un modèle souvent utilisé est ainsi d'effectuer une première introduction au sujet avec une ressource à distance, puis une période en face à face avec un enseignant suivra. Une session de débriefing est souvent aussi ajoutée en fin de formation, quelque temps après celle suivie en face à face.

On notera que l'article "Blended learning" de Wikipedia fournit un autre exemple de scénario didactique « typique » : "A good example of blended learning would be to give a well-structured introductory lesson in the classroom, and then to provide follow-up materials online."

b) L'article « Apprentissage en ligne » (dans sa version actuelle) se contente ensuite d'évoquer, sans les développer, des rubriques qui s'inscrivent sous la notion d'e-learning et pour lesquels le lecteur est renvoyé, en un certain nombre de cas, à d'autres articles de l'encyclopédie en ligne.

## **Plateformes d'apprentissage en ligne**

### **Systemes d'apprentissage**

Voir : [Plate-forme d'apprentissage en ligne](#) dit [CMS](#), [LMS](#), [LCMS](#)

C'est un système informatique de gestion de formation et de gestion de contenu de formation. Le terme anglo-saxon équivalent est L(C)MS pour Learning (Content) Management System.

### **Spécifications**

SCORM (Sharable Content Object Reference Metadata) est une suite de spécifications techniques qui permet aux systèmes d'apprentissage en ligne de trouver, importer, partager, réutiliser, et exporter les [contenus d'apprentissage](#), de manière [normalisée](#).

### **Notions liées**

- \* des concepts de réseaux (principalement basé sur internet)
- \* d'enseignement à distance / [FOAD](#)
- \* d'[auto-apprentissage](#) (très proche de l'[autodidaxie](#), mais avec un support de cours)
- \* de [travail collaboratif](#)
- \* de systématisation

► Il semble qu'une part essentielle du e-learning existant aujourd'hui se fasse autour d'un noyau constitué d'une *plate-forme d'apprentissage en ligne* (comme *Claroline* ou *Dokeos*), « site web qui héberge du contenu didactique et facilite la mise en œuvre de stratégies pédagogiques », précise l'article « Plate-forme d'apprentissage en ligne » de Wikipédia. Ce sont là des pratiques naissantes, à la confluence de l'emploi de TIC en eux-mêmes révolutionnaires et d'une exploitation *didactique* souvent naïve et fréquemment non pensée, ou du moins mal pensée, et que je me propose donc d'interroger.

*That's all, folks!*